International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol.8, No.10 (2015), pp.121-128
http://dx.doi.org/10.14257/ijsip.2015.8.10.14

Fuzzy Time Series Prediction Model and Application based on Fuzzy Inverse

Kun Zhang[1,2], Zhuang Li[1,2*], Hai-feng Wang[1,2] and Hong-xu Wang[1,2]
[1] *College of Computer Engineering*
Qiongzhou University
[2] *Sanya Key Laboratory of Computer Vision*
Qiongzhou University
zk0588@163.com, hitlz@163.com

Abstract

This paper presents a forecast model of fuzzy time series prediction to study the prediction of enrollment. It takes the percentage of enrollment changes as the domain, constructs the inverse fuzzy number and predicts the enrollment of Qiong Zhou University from 2005 to 2013. Compared with the existing models, the mean square error and prediction error of the ameliorated model are smaller and the precision is higher. The authors get the better method to solve the prediction problem based on the inverse fuzzy prediction model.

Keywords: *Percentage, Discrete Domain, Inverse Fuzzy Number, Fuzzy Time Series, Prediction Model*

1. Introduction

In order to solve the problems of fuzziness in natural and social science, there has been many researches on the problem of uncertainty theory, for example: Fuzzy Set Theory, Rough Set Theory, and Evidence Theory, etc. Especially the fuzzy set [1] which created by Zadeh in 1965 was widely used and achieved many results. Many scholars have applied fuzzy set theory to study the time series prediction problems. In 1993, Song and Chissom [2] proposed the first a fuzzy time series forecasting model, after that many scholars put forward different models. These models are used to predict functional new enrollment prediction [2-7], temperature prediction [8], stock index prediction [9], exchange rate forecast [10], data mining [11], road traffic accident forecast, and so on. There are more and more application fields. In the prediction problem of reference [2-7] studied the enrollment of the Alabama University, the average prediction error rate of AFER and the mean square error of MSE is different under different models. This paper improved the prediction model mainly based on the model which is proposed by Saxena, and Sharma and Easo [5]. Literature [5-7] put forward and used the concept "Inverse Fuzzy Number", this paper also obtains a new fuzzy time series forecasting model by this concept. The prediction problem is applied to the QiongZhou University 2004~2013 enrollment, which makes the AFER and MSE are small.

In section 2, this paper simply introduces the basic concept; In section 3, the authors give the basic steps of a new fuzzy time series; In section 4, the prediction of the new model is applied to the QiongZhou University 2004~2013 enrollment, described the prediction process in detail; In section 5, there is the discussion about the same problem applied by Saxena Sharma and the Easo [5], Section 6 is the conclusion.

2. Basic Concepts

Brief Introduction of the basic concepts used by Saxena, Sharma and Easo [5].

ISSN: 2005-4254 IJSIP

Definition 2.1. fuzzy subset X of Nonempty discrete set $P=\{P_1,P_2,\cdots,P_n\}$ is a map given by discourse domain P to closed interval [0, 1]. $\mu_X:P\to[0,1]$. μ_X called prediction the subordinate function of the fuzzy subset X. $\mu_X(P_i),(P_i\in P,i=1,2,\cdots,n)$ called the membership of element P_i to fuzzy set X. Fuzzy set X Can be written as

$$X=\mu_X(P_1)/P_1+\mu_X(P_2)/P_2+\cdots\mu_X(P_n)/P_n.$$

Definition 2.2. According to the sequence of time sequence data points called time series

Definition 2.3. Time series of imprecise data are considered as fuzzy time series

Definition 2.4. The percentage changes in real data defined by history for

$$\frac{e_i-e_{\text{i-1}}}{e_{\text{i-1}}}\times100\%, \tag{1}$$

e_i,e_{i-1} is the real data of history in i , $i-1$year.

Definition 2.5. The definition of prediction error is

$$e_i-f_i \tag{2}$$

Forecast variance is

$$\left(e_i-f_i\right)^2 \tag{3}$$

MSE(Mean Square Error) is

$$MSE=\frac{1}{n}\sum_{i=1}^{n}(e_i-f_i)^2 \tag{4}$$

e_i and f_i is the true history data and forecast data in i years.

Definition 2.6. Forecast error rate is

$$\frac{\left|e_i-f_i\right|}{e_i} \tag{5}$$

AFER（Average Forecasting Error Rate）is

$$AFER=\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\left|e_i-f_i\right|}{e_i}\right)\times100\% \tag{6}$$

e_i and f_i is the true history data and forecast data in i years.

3. New Fuzzy Time Series Forecasting Model

This paper improved Saxena, and Sharma and Easo [5] model, "a fuzzy time series forecasting model of inverse based on fuzzy number”, the basic steps are as follows:

Step 1 giving the real historical data;

Step 2 collating the historical data;

Step 3 constructing the discrete domain;

Step 4 establishing fuzzy inverse formula and the prediction formula;

Step 5 calculating the predicted data of historical data;

Step 6 comparing with other models.

4. The Prediction of QiongZhou University History Enrollment

4.1 Giving the Real Historical Data

QiongZhou University historical enrollment in2004~2013 is shown in Table 1.

Table 1. The Enrollment in QiongZhou University in 2004~2013

Year	Enrollments e_i	Percentage	Year	Enrollments e_i	Percentage
2004	2237		2009	3809	$P_{2009}^{9}=28.987\%$
2005	2852	$P_{2005}^{8}=27.492\%$	2010	4509	$P_{2010}^{7}=18.378\%$
2006	2876	$P_{2006}^{4}=0.842\%$	2011	4410	$P_{2011}^{2}=-2.196\%$
2007	2871	$P_{2007}^{3}=-0.174\%$	2012	4583	$P_{2012}^{6}=3.923\%$
2008	2953	$P_{2008}^{5}=2.856\%$	2013	4397	$P_{2013}^{1}=-4.058\%$

4.2 Collating the Historical Data

Application of formula (1) calculates the percent year to year changes of real historical data. For example:

$$2004\text{-}2005: \quad P_{2005}=\frac{e_{2005}-e_{2004}}{e_{2004}}\times 100\%=\frac{2852-2237}{2237}\times 100\%=27.492\%;$$

$$2010\text{-}2011: \quad P_{2011}=\frac{e_{2011}-e_{2010}}{e_{2010}}\times 100\%=\frac{4410-4509}{4509}\times 100\%=-2.196\%.$$

4.3 Constructing the Discrete Domain

The percentage change of the QiongZhou University 2004~2013 Year Enrollment in the minimum and maximum values is $P_{\min}=-4.058, P_{\max}=28.987$. Use percentage of enrollments of QiongZhou University year to year as data elements to establish the discrete domain:

$$P=\{P_{2013}^{1}=-4.058, P_{2011}^{2}=-2.197, P_{2007}^{3}=-0.174, P_{2006}^{4}=0.842,$$
$$P_{2008}^{5}=2.856, P_{2012}^{6}=3.923, P_{2010}^{7}=18.378, P_{2005}^{8}=27.492, P_{2009}^{9}=28.987\}.$$

4.4 Establishing Fuzzy Inverse Formula and the Prediction Formula

Establishing fuzzy subset on discourse domain P:

$$X_1=\frac{1}{P_{2013}^{1}}+\frac{0.05}{P_{2011}^{2}}+\frac{0}{P_{2007}^{3}}+\cdots+\frac{0}{P_{2009}^{9}},$$
$$X_i=\frac{0}{P_{2013}^{1}}+\cdots+\frac{0.05}{P_{x}^{i-1}}+\frac{1}{P_{y}^{i}}+\frac{0.05}{P_{z}^{i+1}}+\cdots\frac{0}{P_{2009}^{9}}, 2\le i\le 8,$$
$$X_9=\frac{0}{P_{2013}^{1}}+\cdots+\frac{0}{P_{2010}^{7}}+\frac{0.05}{P_{2005}^{8}}+\frac{1}{P_{2009}^{9}}.$$

The Fuzzy subset on discrete domain P which element is a discrete domain of real numbers, Therefore the fuzzy subset can also be called a fuzzy number. The fuzzy numbers above can also give them an operation as follows:

$$J_1=1\div P_{2013}^{1}+0.05\div P_{2011}^{2}+0\div P_{2007}^{3}+\cdots+0\div P_{2009}^{9}=1\div P_{2013}^{1}+0.05\div P_{2011}^{2},$$
$$J_i=0.05\div P_{x}^{i-1}+1\div P_{y}^{i}+0.05\div P_{z}^{i+1}, 2\le i\le 8,$$
$$J_9=0.05\div P_{2005}^{8}+1\div P_{2009}^{9}.$$

There is also provided: $0\div P_{x}^{i}=0, (i=1,2,\cdots,9; x=2005,2006,\cdots,2013)$ So in the operations of fuzzy numbers, fuzzy numbers is actually a real number. For example:

$J_1 = 1 \div P_{2013}^1 + 0.05 \div P_{2011}^2 = 1 \div (-4.058) + 0.05 \div (-2.197) = -0.246427 - 0.022758 = -0.269185.$

Inverse the fuzzy number, can get the inverse fuzzy number formula on discourse domain P.

$$I_i^y = \begin{cases} \dfrac{1+0.05}{\dfrac{1}{P_{2013}^1} + \dfrac{0.05}{P_{2011}^2}}, \\ \dfrac{0.05+1+0.05}{\dfrac{0.05}{P_x^{i-1}} + \dfrac{1}{P_y^i} + \dfrac{0.05}{P_z^{i+1}}}, & 2 \le i \le 8, \\ \dfrac{0.05+1}{\dfrac{0.05}{P_{2005}^8} + \dfrac{1}{P_{2009}^9}}. \end{cases} \tag{7}$$

The same operation of Inverse fuzzy number is as follows:

$$I_1^{2013} = (1+0.05) \div (1 \div P_{2013}^1 + 0.05 \div P_{2011}^2),$$

$$I_i^y = (0.05+1+0.05) \div (0.05 \div P_x^{i-1} + 1 \div P_y^i + 0.05 \div P_z^{i+1}), 2 \le i \le 8,$$

$$I_9^{2009} = (0.05+1) \div (0.05 \div P_{2005}^8 + 1 \div P_{2009}^9).$$

Thus $I_i^y, i = 1, 2, \cdots, 9; y = 2005, 2006, \cdots, 2013$ means the inverse fuzzy number is a real number, y is the year which membership grade is 1. Besides, $\alpha \div 0 = 0.$ when the percentage of enrollment is 0 so the fraction of $\alpha \div 0$ is 0. In the inverse fuzzy number, the inverse fuzzy number is actually a real number. For example:

$$I_6^{2012} = \frac{0.05+1+0.05}{\dfrac{0.05}{p_{2008}^5} + \dfrac{1}{p_{2012}^6} + \dfrac{0.05}{p_{2010}^7}} = \frac{1.1}{\dfrac{0.05}{2.856} + \dfrac{1}{3.923} + \dfrac{0.05}{18.378}}$$

$$= \frac{1.1}{0.017507 + 0.254907 + 0.002721} = \frac{1.1}{0.275135} = 3.998037.$$

4.5 Establishing the Prediction Formula

The predicted data is

$$f_i = e_{i-1} \times (1 + I_i^y \%). \tag{8}$$

f_i is the predicted data of i years, e_{i-1} is the real history data of $i-1$.

4.6 Calculating the Predicted Data of Historical Data

Calculate 2005~2013 prediction data of each year by the inverse fuzzy number formula (7) and a prediction formula (8).

$$I_1 = \frac{1+0.05}{\dfrac{1}{p_{2013}^1} + \dfrac{0.05}{p_{2011}^2}} = \frac{1.05}{\dfrac{1}{-4.058} + \dfrac{0.05}{-2.197}} = \frac{1.05}{-0.246427 - 0.022758}$$

$$= \frac{1.05}{-0.269185} = -3.900663$$

$$f_{2013} = e_{2012} \times (1 + I_1\%) = e_{2012} \times (1 - 0.03900663) = 4583 \times 0.960993 = 4404$$

$$I_5 = \frac{0.05+1+0.05}{\frac{0.05}{p_{2006}^4}+\frac{1}{p_{2008}^5}+\frac{0.05}{p_{2012}^6}} = \frac{1.1}{\frac{0.05}{0.842}+\frac{1}{2.856}+\frac{0.05}{3.923}}$$

$$= \frac{1.1}{0.059382+0.350140+0.012745} = \frac{1.1}{0.422267} = 2.604987$$

$$f_{2008} = e_{2007} \times (1+I_5\%) = e_{2007} \times (1+2.604987\%) = 2871 \times 1.026050 = 2946$$

$$I_9^{2009} = \frac{0.05+1}{\frac{0.05}{p_{2005}^8}+\frac{1}{p_{2009}^9}} = \frac{1.05}{\frac{0.05}{27.492}+\frac{1}{28.987}} = \frac{1.05}{0.001819+0.034498} = 28.912080$$

$$f_{2009} = e_{2008} \times (1+I_9^{2009}\%) = e_{2008} \times (1+28.912080\%) = 2953 \times 1.289121 = 3807$$

Table 2. Value Prediction of QiongZhou University in 1994~2013

Year	Enrollments e_i	Predicted Value f_i	Prediction Error	Erediction Square Error	$\frac{\lvert e_i - f_i \rvert}{f_i}$
2004	2237				
2005	2852	2840	12	144	0.004208
2006	2876	2886	-10	49	0.003477
2007	2871	2870	1	1	0.000348
2008	2953	2946	7	49	0.002370
2009	3809	3807	2	4	0.000525
2010	4509	4416	93	8649	0.020625
2011	4410	4439	-29	841	0.006576
2012	4583	4586	-3	9	0.000655
2013	4397	4404	-7	49	0.001592
AFER					0.4486%
MSE				1088	

Calculate the prediction data of each year, fill in the table 2. Calculate the forecasting error and variance by using formula (2) and (3); Calculate prediction error rate by formula (5). Finally calculate the mean square error of MSE and the average prediction error rate of AFER by formula (4) and (6) then fill in Table 2

5. Comparing with the Existing Best Method

Saxena, and Sharma and Easo [5] (published in 2012) points out the calculating data: the forecast data of Alabama University enrollment in 1978~1992 get the good results AFER=0.34% MSE=9167. The model "provides the minimum AFER and MSE" [5], the proposed model is better than the existing model (proposed model before 2012) produced the better accuracy. The follows are the QiongZhou University enrollment in 2004~2013 studied by the application of the model of Saxena, Sharma and Easo [5]. The basic steps of the model are:

The first step: To determine the domain interval according to the historical data by newborn registration number of the changes, and split it into 7 equal length sub interval.

The second step: To count the frequency percentage statistics enrollment changes. To give each sub interval the smaller interval smaller interval according to the frequency of each sub interval and give the language value to the fuzzy sub, then get the midpoint of each fuzzy sub interval.

The third step: To make predictions by using the prediction formula:

$$t_j = \begin{cases} \dfrac{1+0.5}{\dfrac{1}{a_1}+\dfrac{0.5}{a_2}}, & if \quad j=1, \\ \dfrac{0.5+1+0.5}{\dfrac{0.5}{a_{j-1}}+\dfrac{1}{a_j}+\dfrac{0.5}{a_{j+1}}}, & if \quad 2 \le j \le n-1, \\ \dfrac{0.5+1}{\dfrac{0.5}{a_{n-1}}+\dfrac{1}{a_n}}, & if \quad j=n. \end{cases} \tag{9}$$

The midpoints of Fuzzy sub interval X_{j-1}, X_j, X_{j+1} are expressed by a_{j-1}, a_j, a_{j+1}. But parameter t_j is expressed as a percentage of enrollment prediction. The percentage change of QiongZhou University Enrollment in 2004~2013, the minimum and maximum values are $P_{\min} = -4.058, P_{\max} = 28.987$. Apply the model of Saxena, and Sharma and Easo [5]. Domain prediction problem is the enrollment of QiongZhou University in 2004~2013. The final segmentation of fuzzy sub interval is shown in Table 3

Apply formula (9) to calculate the forecast data of QiongZhou University enrollment 2004~2013 as shown in Table 4.

Table 3. Sub Fuzzy Interval and Its Points

Fuzzy Sub Interval	Percentage inside	Midpoint
X_1=[-4.10,-2.52]	$P_{2013}^{X_1} = -4.058$	$a_1 = -3.31$
X_2=[-2.52,-0.94]	$P_{2011}^{X_2} = -2.197$	$a_2 = -1.73$
X_3=[-0.94,0.64]	$P_{2007}^{X_3} = -0.174$	$a_3 = -0.15$
X_4=[0.64,2.22]	$P_{2006}^{X_4} = 0.842$	$a_4 = 1.43$
X_5=[2.22,3.80]	$P_{2008}^{X_5} = 2.856$	$a_5 = 3.01$
X_6=[3.80,5.38]	$P_{2012}^{X_6} = 3.923$	$a_6 = 4.59$
X_7=[5.38,10.12]		$a_7 = 6.85$
X_8=[10.12,14.86]		$a_8 = 12.49$
X_9=[14.86,19.60]	$P_{2010}^{X_9} = 18.378$	$a_9 = 17.23$
X_{10}=[19.60,24.34]		$a_{10} = 21.79$
X_{11}=[24.34,26.71]		$a_{11} = 25.53$
X_{12}=[26.71,27.90]	$P_{2005}^{X_{12}} = 27.492$	$a_{12} = 27.31$
X_{13}=[27.90,29.80]	$P_{2009}^{X_{13}} = 28.987$	$a_{13} = 28.85$

Table 4. The Predicted Value of QiongZhou University Enrollment in Another Model

Year	Enrollments E_i	Predicted Value F_i	Prediction Error	Prediction Square Error	$\frac{\lvert E_i - F_i \rvert}{E_i}$
2004	2237				
2005	2852	2845	7	49	0.002454
2006	2876	2829	47	2209	0.016342
2007	2871	2781	90	8100	0.031348
2008	2953	2944	9	81	0.003048
2009	3809	3789	20	400	0.005251
2010	4509	4439	70	4900	0.015525
2011	4410	4487	-77	5929	0.017460
2012	4583	4603	-20	400	0.004364
2013	4397	4467	-70	4900	0.015920
AFER					1.2412%
MSE				2997	

From Tables 2 and 4 we can know that for the prediction problem of QiongZhou University enrollment in 2004~2013, AFER=0.4486%, MSE=1088 in this model. AFER=1.2412%, MSE=2997 in Saxena, and Sharma and Easo's model. This model is better than the model of Saxena, and Sharma and Easo [5], the predictive accuracy of the model is better.

6. Conclusion

This model is a modified model of the model of Saxena, Sharma and Easo [5]. This study improves the domain selection and the establishment of inverse fuzzy number. The method is simpler, the predictive accuracy of the model is higher. It provides a new method to solve the problem of how to improve the prediction accuracy and predict unknown data of known data in the time series prediction problems.

Acknowledgements

This work is partially supported by the 2014 National Students' Innovation and Entrepreneurship Training Program (No.201411100061, 201411100066, 201411100068, 201411100069); the Hainan applied technology development and demonstration of special project (No.zdxm2014087); the 2013 Hainan education science research topics of the 12th Five-Year Plan (No.QJY125119); The Hainan Planning projects of philosophy and Social Sciences (No.HNSK14-50); the scientific and technological cooperative project for college and region of Sanya (No.2014YD14, 2014YD24, 2014YD30, 2014YD39); the Agricultural Science and Technology Innovation Project of Sanya (No.2014NK26); the key Laboratory of Sanya Project(No.L1410).

References

[1] L. A. Zadeh, "Fuzzy sets", Information and Control, vol. 8, no. 3, **(1965)**, pp. 338-353.
[2] Q. Song and B. S. Chissom, "Forecasting enrollments with fuzzy time series—part I", Fuzzy Set and Systems, vol. 54, **(1993)**, pp. 1-9.
[3] Q. Song and B. S. Chissom, "Fuzzy time series and its models", Fuzzy Set and Systems, vol. 54, **(1993)**, pp. 169-277.
[4] Q. Song and B. S. Chissom, "Forecasting enrollments with fuzzy time series—part II", Fuzzy Set and Systems, vol. 62, **(1994)**, pp. 1-8.

[5] P. Saxena, K. Sharma and S. Easo, "Foreeca enrollment based on fuzzy time series with higher forecast accuracy rate", International Journal of Computer Technology and Applications, vol. 3, no. 3, **(2012)**, pp. 957-961.

[6] T. A. Jilani, S. M. A. Burney, and C. Ardil, "Fuzzy metric approach for fuzzy time series forecasting based on frequency density based partitioning", World Academy of Science, Engineering and Technology, vol. 34, **(2007)**, pp. 1-6.

[7] M. Stevenson and J. E. Porter, "Fuzzy time series forecasting using percentage change as the universe of discourse", World Academy of Science, Engineering and Technology, vol. 55, **(2009)**, pp. 154-157.

[8] N. Y. Wang and S. M. Cheng, "Temperature prediction and TAIFEX forecastion based on automatic clustering techniques and two-factors high-order fuzzy time series", Expert Systems with Applications, vol. 36, no. 2,**(2009)**, pp. 2143-2154.

[9] H. H. Chu, T. L. Chen, C. H. Cheng, C. C. Huang, "Fuzzy dual-factor time-service for stock index forecasting", Expert Systems with Application, vol. 36, **(2009)**, pp. 165-171.

[10] Y. H. Leu, C. P. Lee and Y. Z. Jou, "A distance-based fuzzy time series model for exchange rates forecasting. Expert Systems with Applications", vol. 36, no. 4, **(2009)**, pp. 8107-8114.

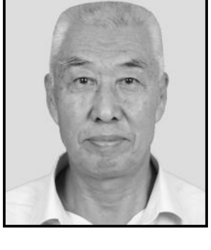
[11] C. H. Chen, T. P. Hong and V. S. Tseng, "Fuzzy data mining for time-series data", Applied Soft Computing, vol. 12, **(2012)**, pp. 536-542.

Authors

Kun Zhang, he was born in 1981, received the B.S. degree in 2004 from Harbin Normal University, Harbin, PR China, the M.S. degrees in School of Computer Science, Maritime Affairs University of Dalian, Dalian PR China in 2011. He is professor with College of Computer Engineering, Qiongzhou University, Sanya Hainan PR China. His research interests include Embedded System Design, Database design, Intelligent Data Analysis and Data Mining.

Haifeng Wang, he was born in 1980, received the M.S. degree in college of information and computer Engineering from northeast forestry university in 2009.He is a lecturer with College of Electronics and Information Engineering, Qiongzhou University, Sanya, PR China. Since 2012 he has been pursuing his PhD degree in northeast forestry university. His research interests include computer vision, image segmentation, pattern recognition, forestry intelligent equipment and Information control

Zhuang Li, he was born in 1967, received the PhD degree in 2007 from Haerbin Institute of Technology, Harbin, PR China. He is professor with College of Computer Engineering, Qiongzhou University, Sanya Hainan PR China. His research interests include Inverse Calculation, Intelligent Data Analysis and Data Mining.



Hongxu Wang, he was born in 1946, received the B.S. degree in 1987 from the Department of Mathematics of The Jilin University, Changchun, PR China. He is professor with Qiongzhou University, Sanya Hainan PR China. His research interests include Fuzzy control, information processing.